2021-S-0004 Standard Practices for Evaluating Measurement Uncertainty of Quantitative Measurements in Forensic Toxicology

Forensic Toxicology Subcommittee
Toxicology Scientific Area Committee
Organization of Scientific Area Committees (OSAC) for Forensic Science

Draft OSAC Proposed Standard

2021-S-0004 Standard Practices for Evaluating Measurement Uncertainty of Quantitative Measurements in Forensic Toxicology

Prepared by
Forensic Toxicology Subcommittee
Version: 1.0
2020

Disclaimer:

This OSAC Proposed Standard was written by the Forensic Toxicology Subcommittee/ Toxicology Scientific Area Committee of the Organization of Scientific Area Committees (OSAC) for Forensic Science following a process that includes an open comment period. This Proposed Standard will be submitted to a standards developing organization and is subject to change.

There may be references in an OSAC Proposed Standard to other publications under development by OSAC. The information in the Proposed Standard, and underlying concepts and methodologies, may be used by the forensic-science community before the completion of such companion publications.

Any identification of commercial equipment, instruments, or materials in the Proposed Standard is not a recommendation or endorsement by the U.S. Government and does not imply that the equipment, instruments, or materials are necessarily the best available for the purpose.

To be placed on the OSAC Registry, certain types of standards first must be reviewed by a Scientific and Technical Review Panel (STRP). The STRP process is vital to OSAC's mission of generating and recognizing scientifically sound standards for producing and interpreting forensic

science results. The STRP shall provide critical and knowledgeable reviews of draft standards or of proposed revisions of standards previously published by standards developing organizations (SDOs) to ensure that the published methods that practitioners employ are scientifically valid, and the resulting claims are trustworthy.

The STRP panel will consist of an independent and diverse panel, including subject matter experts, human factors scientists, quality assurance personnel, and legal experts, which will be tasked with evaluating the proposed standard based on a comprehensive list of science-based criteria.

For more information about this important process, please visit our website at: https://www.nist.gov/topics/organization-scientific-area-committees-forensic-science/scientific-technical-review-panels

Standard Practices for Evaluating Measurement Uncertainty of Quantitative Measurements in Forensic Toxicology

Foreward

This document was developed to provide the minimum requirements for evaluating measurement uncertainty for quantitative measurements in forensic toxicology laboratories. Measurement uncertainty is required to ensure confidence, reliability, and proper interpretation of test or calibration results. It is also one of the components used to establish measurement traceability. This standard was developed by the Toxicology Subcommittee of the Organizational Scientific Area Committee.

45 **Table of Contents**

46

57
58

1 Scope

This document provides minimum requirements for evaluating measurement uncertainty for quantitative results in forensic toxicology. The document is for testing activities and calibration of breath alcohol measuring instruments and provides direction on evaluation of components, bias, calculations, and reporting. It does not address evaluating measurement uncertainty for breath alcohol testing, this topic will be covered in a different document.

2 Normative References

National Institute of Standards and Technology, SOP 29-Standard Operating Procedure for the Assignment of Uncertainty (February 2018). Available for download at https://www.nist.gov/pml/weights-and-measures/laboratory-metrology/standard-operating-procedures

Joint Committee for Guides in Metrology (JCGM) Evaluation of Measurement Data-Guide to the Expression of Uncertainty in Measurement (GUM) (GUM 1995 with minor corrections) (Sevres, France: International Bureau of Weights and Measures [BIPM]-JCGM 100], September, 2008. Available at https://www.bipm.org/en/publications/guides/gum.html

SLR Ellison and A Williams (Eds). Eurachem/CITAC Guide: Quantifying Uncertainty in Analytical Measurement, Third edition, (QUAM: 2012 P1) Available for download at http://www.eurachem.org/index.php/publications/guides

Joint Committee for Guides in Metrology (JCGM), *International vocabulary of metrology – Basic and general concepts and associated terms* (**VIM**), 3rd ed. (Sèvres, France: International Bureau of Weights and Measures [BIPM]-JCGM 200, 2012) (2008 with minor corrections). Available for download at http://www.bipm.org/en/publications/guides/vim.html

ANSI/ASB Standard 017, Standard Practices for Measurement Traceability in Forensic Toxicology, First Edition, 2018. Available for download at https://asb.aafs.org/wp-content/uploads/2018/06/017_Std_e1.pdf

3 Terms and Definitions

For purposes of this document, the following definitions and acronyms apply.

3.1
accuracy
Closeness of agreement between a measured quantity value and a true quantity value of a measurement.

3.2
analytical run (batch)
A set of standards, controls, and/or case samples that are contemporaneously prepared and/or
analyzed in a particular sequence.

3.3
bias, statistical
A systematic tendency for estimates or measurements to be above or below their true values.

Note 1: Statistical bias arises from systematic as opposed to random error.

Note 2: Statistical bias can occur in the absence of prejudice, partiality, or discriminatory intent.

3.4
calibration
Operation that, under specified conditions, establishes a relationship between the quantity value and corresponding indications.

3.5
calibrator
Measurement standard used in calibration.

3.6
certified reference material
CRM
Reference material (RM) characterized by a metrologically valid procedure for one or more specified properties, accompanied by a certificate that provides the value of the specified property, its associated uncertainty, and a statement of metrological traceability.

3.7
control
Material of known composition that is analyzed along with unknown samples(s) in order to evaluate the performance of an analytical procedure.

3.8
limit of detection
LOD
An estimate of the lowest concentration of an analyte in a sample that can be reliably differentiated from blank matrix and identified by the analytical method.

3.9
lower limit of quantitation
LLOQ
An estimate of the lowest concentration of an analyte in a sample that can be reliably measured with acceptable bias and precision.

3.10
measurand
The quantity intended to be measured.

3.11
measurement traceability
(metrological traceability)
Property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.

3.12
precision
The measure of the closeness of agreement between a series of measurements obtained by replicate measurements on the same or similar samples.

3.13
repeatability
Measurement precision under a set of conditions that includes the same measurement procedure, same operators, same measuring system, same operating conditions, same conditions and same location, and replicate measurements on the same or similar objects over a short period of time.

3.14
reproducibility
Measurement precision under a set of conditions that includes different locations, operators, measuring system, and replicate measurements on the same or similar objects.

4 Measurement Uncertainty

4.1 Background

Quantitative values obtained from measurement processes have an expected variability. Repeated measurements will result in different values each time a measurement is made provided the measuring system has sufficient resolution to allow those differences to be seen. Each time a measurement is made, the measured value depends on numerous factors including setup and capability of the measuring system, the exact measurement method (procedure), and the person performing the measurement.

Measurement Uncertainty (MU) is an estimate of the potential variability of a measurement based on the information known about the measurand and the measurement method. The measurement may be part of the test, a calibration method, or the final reported test or calibration result. "Measurement uncertainty does not imply doubt about the validity of a measurement; on the contrary, knowledge of the uncertainty implies increased confidence in the validity of the measurement result[1]"

Laboratory stakeholders require tests and calibrations performed to be reliable, accurate, and comparable. MU is an important parameter describing the confidence, as well as limitations, of measurement results. Comparison of quantitative test or calibration results between laboratories or evaluation of quantitative results in relation to a legal specification or requirement necessitates knowledge of the MU.

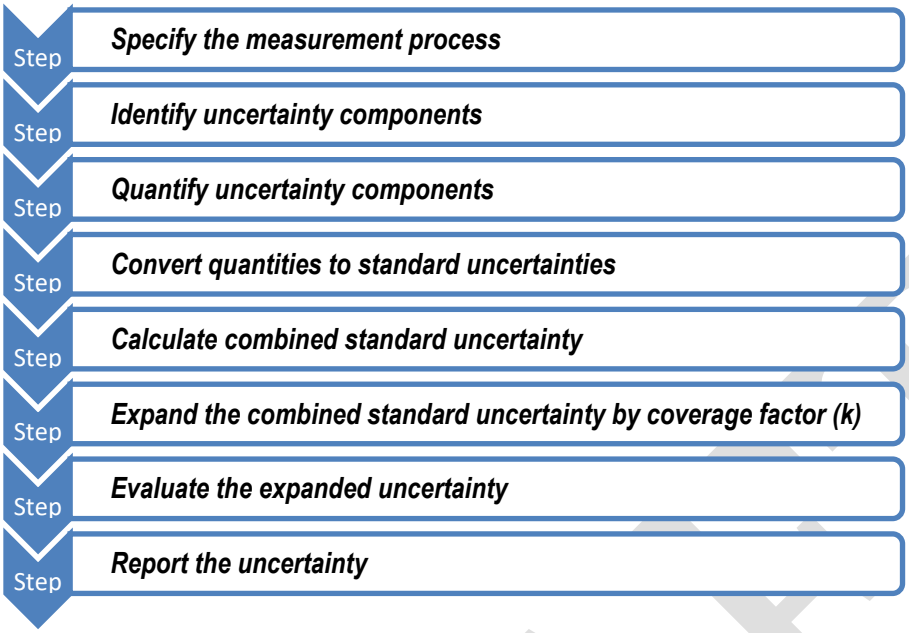
The National Institute for Standards and Technology (NIST) has developed an 8-step process for evaluating and reporting MU (Figure 1).[2] This framework established by NIST conforms to the principles set forth in the Joint Committee for Guides in Metrology (JCGM) Evaluation of Measurement Data-Guide to the Expression of Uncertainty in Measurement (GUM[3]) and is a helpful reference.

[1] SLR Ellison and A Williams (Eds). Eurachem/CITAC Guide: Quantifying Uncertainty in Analytical Measurement, Third edition, (QUAM: 2012 P1) Available for download at http://www.eurachem.org/index.php/publications/guides

[2] National Institute of Standards and Technology, SOP 29-Standard Operating Procedure for the Assignment of Uncertainty (April 2021). Available for download at https://www.nist.gov/system/files/documents/2019/05/13/sop-29-assignment-of-uncertainty-20190506.pdf

[3] Joint Committee For Guides in Metrology (JCGM) Evaluation of Measurement Data-Guide to the Expression of Uncertainty in Measurement (GUM) (GUM 1995 with minor corrections) (Sevres, France: International Bureau of Weights and Measures [BIPM]-JCGM 100], September, 2008. Available at http://bipm.org/en/publications/guides/gum.html

Figure 1: The NIST 8-Step Process for Evaluating and Reporting Measurement Uncertainty[4]



4.2 Requirements for Measurement Uncertainty for Quantitative Determinations

4.2.1 General Requirements

4.2.1.1 Laboratories shall have and apply procedures for evaluating MU for methods used to calibrate breath alcohol instruments and for test methods that produce a quantitative test result.

4.2.1.2 MU is specific to each measurement process and shall be evaluated separately for each analyte in each testing or calibration method.

In testing, this requires that each combination of analyte, extraction and analytical technique be evaluated separately. Multiple matrices may have to be evaluated separately based on results of method validation.

4.2.1.3 Using the largest evaluated MU for more than one analyte within a method or one analyte across methods is not acceptable.

4.2.1.4 Test and Calibration Methods for which the MU is evaluated shall meet the minimum requirements set forth in:

[4] Adapted from ASCLD/LAB Guidance on the Estimation of Measurement Uncertainty-Annex D
Note: Document can be obtained from anab@anab.org

a. *ANSI/ASB Standard 017,* Standard Practices for Measurement Traceability in Forensic Toxicology.
b. *ANSI/ASB Standard 036,* Standard Practices for Method Validation in Forensic Toxicology

4.2.2 Step 1: Specify the Measurement Process

The measurandshall be defined and documented. This can be in the form of a written statement, a visual diagram, and/or a mathematical expression.

NOTE: To be clear about the measurement process for which the MU evaluation is for, it is important to be as specific as possible when defining the measurand. To distinguish one measurement process from another within a laboratory, it may be necessary to include a reference to a specific type of equipment used or a specific procedure in the statement defining the measurand.

EXAMPLES:

Testing of biological samples
Concentration of ethanol (g/100mL) in ante-mortem whole blood
Concentration of oxycodone (mg/kg) in a sample of liver homogenate

Calibration of breath alcohol measuring instruments
Calibration of XYZ model breath alcohol measuring instrument using dry gas certified reference material

4.2.3 Step 2: Identify Uncertainty Components

Minimum method components that shall be considered, as applicable, in an evaluation of MU include:

a. Certified reference material(s) and calibrations of equipment used to establish measurement traceability;

b. Data from the measurement process (i.e. repeatability, reproducibility or from intermediate measurement conditions)

c. Human factors (e.g., multiple analysts performing the same measurement method, experience, training , etc.);

d. Sampling conducted during the measurement method;

e. Sample preparation; and

f. Environmental conditions during the measurement process.

4.2.4 Step 3: Quantify Uncertainty Components

Uncertainty components shall be quantified. All digits shall be carried through calculations until final expanded measurement uncertainty is determined. Only then should rounding and significant figure rules be applied.

The GUM (2.3.2) refers to the method of evaluation of one or more uncertainty components as:

> *Type A* evaluation (of uncertainty): method of evaluation of uncertainty by the statistical analysis of series of observations (e.g., relative standard deviation of a historical data set of quality control results)
>
> *Type B* evaluation (of uncertainty): method of evaluation of uncertainty by means other than the statistical analysis of series of observations (e.g., obtaining the uncertainty associated with a CRM from its certificate of analysis)

The method of evaluation, Type A or Type B, will be determined for each component identified. It is most common to have a mixture of the two methods where some identified uncertainty components are quantified using a Type A method of evaluation and some identified uncertainty components are quantified using a Type B method of evaluation.

Any double-counting of a component will result in an overestimation of the measurement uncertainty and should be avoided, when possible. However, overestimation is generally more desirable than underestimation.

A record shall be maintained for *Type A* and *Type B* evaluations.

4.2.4.1 Minimum requirement(s) for data used in *Type A* evaluations:

4.2.4.1.1 Shall come from method validation and/or ongoing quality control (measurement assurance program) for the measurement method.

a. Method validation may include the evaluation of one or more specific uncertainty components.

b. Data from proficiency tests may only be used if the proficiency test has established metrological traceability for the quantitative value of the proficiency test. A consensus value does not establish metrological traceability.

4.2.4.1.2 Shall be representative of the measurand that will be tested or calibrated.

4.2.4.1.3 Shall be representative of the range (e.g., matrix, or detector response over the expected concentration range, etc.) of the measurements made.

4.2.4.1.4 Shall be evaluated according to the size and distribution of the statistical sample.

4.2.4.2 Establishing a quantity value for *Type A* evaluations

To appropriately evaluate the magnitude of uncertainty for the measurement process using Type A evaluation, calculate a standard deviation or a relative standard deviation using historical data for each identified *Type A* uncertainty component. Typically, method performance is best represented by measurements of quality control (QC) samples taken over multiple instrumental batches, each with different instrument calibrations (cf. ANSI/ASB Standard 036, 2019). A graphical representation of all QC measurements used for the Type A uncertainty component that demonstrates statistical control of the measurements used shall be maintained. Additional methods may also be used to ensure statistical control.

If multiple QC measurements are available in each instrumental batch, all QC measurements can be included when computing the standard deviation or relative standard deviation. Inclusion of multiple QC measurements in the computation of the standard deviation will bias the standard uncertainty estimate slightly if the QC data exhibits any batch-to-batch variation but mitigates the need for more complex standard deviation computations. If needed, other statistical methods, such as the ANOVA method outlined in section 8.2.2.3.4 of ANSI/ASB Standard 036 or random subsampling of the QC data to select a single representative QC measurement from each batch, can be used to correct for this bias.

If the result to be reported for a specimen will be either an individual measured value or the average of multiple measured values from a single instrumental batch, the standard deviation or the relative standard deviation shall be used as the Type A standard uncertainty for the reported specimen value. Setting aside the slight bias produced if the standard deviation is computed from data containing multiple QC measurements in each batch, this standard uncertainty should provide an assessment of the Type A uncertainty that is either on target or conservative (i.e., larger than necessary) for the reported specimen value.

If the the result to be reported for a specimen will be the average of measured values from multiple instrumental batches, the standard deviation or the relative standard deviation divided by the square root of the number of instrumental batches used when averaging the specimen data shall be used as the Type A standard uncertainty for the reported specimen values. Division by the square root of the number of batches converts the standard deviation for single-batch results into the standard deviation of the mean of multiple-batch results. As above, setting aside the slight bias produced if the standard deviation is computed from data containing multiple QC measurements in each batch, this standard uncertainty should provide an assessment of the Type A uncertainty that is either on target or conservative for the reported specimen value.

4.2.4.2.1 Testing Laboratories

4.2.4.2.1.1 Use of Validation Data

Validation data may initially be used for the Type A uncertainty component. Continued use of validation data for this uncertainty component requires that laboratories demonstrate the data is representative of the data generated during day-to-day analysis by analysts who have demonstrated competence.

4.2.4.2.1.2 Multiple controls within the same method

For methods where validation has demonstrated constant variance across the entire calibration range (homoscedasticity) as shown through the use of residual plots for the calibration curve or other statistical means, laboratories shall use either:

a. Combined data from all controls analyzed; or

b. Select data from one specified control (e.g., a control at or near a legal specification).

For methods where validation has demonstrated that variance is not constant across the entire calibration range (heteroscedasticity), laboratories shall establish a procedure for how MU will be calculated. Procedures may include:

a. Utilize the Type A data from the control producing the largest variance; or

b. Perform an in-depth evaluation to determine where the variation changes occur across the calibration range and establish an appropriate uncertainty to report based on where these variation changes occur; or

c. Utilize the Type A data from the control at the concentration closest to the sample concentration. This is acceptable only when an evaluation of the difference in the standard deviation between the two applicable control levels does not impact the evaluation of conformance with a legal specification.

4.2.4.2.1.3 Multiple Analysts/Instruments/Laboratories

For a method that has been validated on multiple instruments or in multiple laboratories by analysts who have demonstrated competence, provided that quality control criteria for acceptance and reporting criteria are the same across all instruments and laboratories, calculate MU using control data in accordance with Section 4.2.4.2.1.4..

4.2.4.2.1.4 Quality Control Data

Appropriate methods for calculating MU using quality control data include, but are not limited to:

a. Calculation of MU using control data generated since validation or first day-to-day use of the method;

b. Calculation of a rolling MU where the laboratory chooses to include a set number of data points from the most recent analyses. The data shall be representative of the performance of the method; or

c. Calculation of a batch-specific MU based on use of data from only the current analytical batch. This method is more commonly used for non-routine analysis where limited data points are available.

4.2.4.2.2 Calibration of Breath Alcohol Measuring Instruments

4.2.4.2.2.1 Use of Validation Data

Validation data may initially be used for the Type A uncertainty component. Continued use of validation data for this uncertainty component requires that laboratories demonstrate that the data is representative of the data generated during day-to-day calibration of breath alcohol measuring instruments by personnel who have demonstrated competence.

4.2.4.2.2.2 Multiple measurement standards within the same method

For methods that have demonstrated constant variance across the entire calibration range as shown through the use of residual plots for the calibration curve or other statistical means, laboratories may either:

a. Combine data from all measurement standards analyzed to estimate a single MU; or

b. Calculate the measurement uncertainty at each measurement standard concentration.

For methods where validation has demonstrated that variance is not constant across the entire calibration range, laboratories may either:

a. Perform an in-depth evaluation to determine where the variance changes occur across the calibration range and establish an appropriate uncertainty to report based on where these variance changes occur; or

b. Calculate the MU at each measurement standard concentration across a population of instruments or for an individual breath alcohol measuring instrument.

4.2.4.2.2.3 Use of Measurement Standard Data or Quality Control Data

Appropriate methods of selecting measurement standard data or quality control data include, but are not limited to, the following across a population of instruments or for an individual instrument:

a. Calculation of MU using measurement standard data generated since validation or first day-to-day use of the instrument; or

b. Calculation of a rolling MU where the laboratory chooses to include a set number of data points. The data shall be representative of the performance of the instrument.

4.2.4.3 Minimum requirements for *Type B* evaluations:

Components requiring a *Type B* evaluation may include: uncertainty associated with a certified reference material, uncertainty of a reference material, and/or uncertainty from equipment calibration (e.g., balance, volumetric flask, pipette, barometer, or thermometer).

4.2.4.3.1 Shall consider all components that are not accounted for in a Type A evaluation.

4.2.4.3.2 Shall account for all identified and significant systematic bias (see 4.2.6.1).

4.2.4.3.3 Shall be handled according to the assumed distribution of the quantity value.

4.2.4.4 Establishing a quantity value for *Type B* evaluations

4.2.4.4.1 For component(s) used in the preparation of a calibrator, the components can be quantified individually or as a group for the calibrator.

a. If estimating uncertainty over the full calibration range, use the largest standard deviation calculated above;

b. If estimating the uncertainty for multiple concentration ranges, use the largest standard deviation calculated above for each concentration range, respectively;

c. If estimating the uncertainty at each calibrator or measurement standard concentration separately, use the value for the applicable calibrator.

If the test or calibration method includes the preparation of multiple calibrators or measurement standards, the individual components can be quantified individually across all calibrator concentrations (e.g. a single component quantity value can be used for the pipette uncertainty that adequately covers the pipettes used to prepare all calibrator concentrations) and then a or b below can be applied. Alternatively, the components can be quantified as a group for each calibrator concentration and then a - c applied.

Depending on the measurement process, these components related to calibrator preparation, typically requiring a *Type B* evaluation, may be accounted for by on-going quality control data (*Type A*).

4.2.5 Step 4: Convert Quantities to Standard Uncertainties

Quantify all uncertainty components as a standard uncertainty of the quantity values and in the same measurement unit or in a measurement unit relative to the quantity values.

4.2.5.1 *Type A* evaluations

Typically, an assessment of Type A uncertainty is calculated to be a standard uncertainty. If not already presented as a standard uncertainty, divide by the appropriate factor (*e.g.,* 2 or 3) to convert to a standard uncertainty.

4.2.5.2 *Type B* evaluations

If not reported by the manufacturer as a standard uncertainty, the appropriate probability density function for the component needs to be used to compute one standard deviation or relative standard deviation associated with the specified distribution.

If reported by the manufacturer as an expanded uncertainty, divide by the appropriate coverage factor (e.g., 2 or 3), to arrive at a standard uncertainty.

4.2.6 Step 5: Calculate the combined standard uncertainty

Calculate the combined standard uncertainty using each uncertainty contributor quantity value. Acceptable methods to do so include the root sum of the squares formula and the Monte Carlo[5] method.

After the combined standard uncertainty is calculated, components may be individually evaluated for significance. A component is deemed significant if it impacts the least significant digit in the reported value for MU.[6] Components determined to be insignificant may be removed from the uncertainty calculations.

Note: If multiple individual components are removed from the uncertainty combination, then the aggregate impact of the removed components should be evaluated.

---

[5] Joint Committee For Guides in Metrology (JCGM) Evaluation of Measurement Data-Guide to the Expression of Uncertainty in Measurement (GUM)-Supplement 1-Propagation of distributions using a Monte Carlo Method (Sevres, France: International Bureau of Weights and Measures [BIPM]-JCGM 101:2008], September, 2008. Available at https://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf

[6] National Institute of Standards and Technology, SOP 29-Standard Operating Procedure for the Assignment of Uncertainty (February 2018). Available for download at https://www.nist.gov/pml/weights-and-measures/laboratory-metrology/standard-operating-procedures

4.2.6.1 Evaluation of bias[7]

Measurement accuracy encompasses both precision and bias. A measurement is more accurate when it has less bias and greater precision. The GUM states "it is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."

An evaluation of bias may not always be possible. An evaluation of bias requires one or more controls prepared with metrological traceability, having a known reference value and uncertainty.

4.2.6.1.1 The general approach to bias evaluation shall:

a. Determine if bias is present by comparing measurement standard or control data to reference values with established metrological traceability.

b. Estimate the combined uncertainty without considering the relevant bias.

c. Compare the bias with the combined standard uncertainty.

   i. Where the bias is less than the combined standard uncertainty, $bias < u_c$, the bias is viewed as not significant and may be neglected or included as a component in the estimation of uncertainty.

   ii. Where the bias is greater than the combined standard uncertainty, $bias > u_c$, it is viewed as significant and additional action is required, see 4.2.6.1.1.1 and 4.2.6.1.1.2.

4.2.6.1.1.1 Testing

a. Eliminate or reduce the bias until it is not significant; or

b. Correct the measurement result for the bias, including the uncertainty of the correction in the evaluation of uncertainty. Both the observed measurement result and the corrected measurement result with the estimation of MU shall be reported; or

c. Report the measurement result and the expanded MU with bias included. The method used to include the bias in the expanded uncertainty shall be a statistically valid method[7] and in compliance with the GUM; or

d. Report the observed measurement result, the MU, and the bias.

[7] Section 3.2.5 of NIST SOP 29 (2019)

4.2.6.1.1.2 Calibration of Breath Alcohol Instruments

Eliminate or reduce bias until it is not significant by repeating the adjustment process and/or performing the appropriate repair.

4.2.7 Step 6: Calculate the expanded uncertainty

4.2.7.1 A coverage factor (*k*) shall be determined using a Student's t-distribution[8] based on the degrees of freedom to provide the desired level of confidence.

4.2.7.2 The minimum coverage probability for all quantitative test results and calibration results shall be 95.45 % (often referred to as approximately 95 %).

4.2.8 Step 7: Evaluate the expanded uncertainty

4.2.8.1 A determination whether the evaluated measurement uncertainty is acceptable shall be made by the laboratory. The laboratory is responsible for supporting their decision. As applicable, minimum aspects to consider include:

a. Stakeholder interests;

b. Legal requirements;

c. The relationship between the reported test or calibration quantitative value and the expanded MU; particular consideration shall be taken around the LLOQ/LOD (e.g., an expanded MU of 0.01 ng/mL for a method with an LLOQ of 0.01 ng/mL should prompt the laboratory to reevaluate the LLOQ.); and

d. The relationship between the quality control limits for the method and the expanded measurement uncertainty (e.g., ± 20 % quality control limits for a method with expanded MU of 10 %. For any single analytical batch, this QC limit would allow a variation of up to 20% which exceeds the stated expanded MU for the method. This should prompt the laboratory to reevaluate the quality control limits to ensure the MU statement will always be correct).

4.2.9 Step 8: Report the expanded uncertainty

4.2.9.1 The estimated MU shall be included in the test/calibration report or an attachment to the report for all quantitative test results in accordance with the *ANSI/ASB Standard 053 Standard for Reporting in Forensic Toxicology* and for all calibrations.

a. The MU shall be reported as an expanded uncertainty and include the coverage probability for testing laboratories

[8] Table G2 of the GUM, another reference table, or appropriate statistical software

b. The MU shall be reported as an expanded uncertainty and include the coverage factor, k, and the coverage probability for calibration laboratories.

c. The measurement result shall include the measured quantity value, y, along with the associated expanded uncertainty, U, and the measurement result should be reported as $y \pm U$ where U is consistent with the units of y. Specific applications may warrant use of a different format than $y \pm U$.

d. The expanded uncertainty should be reported to at most 2 significant figures unless the laboratory has a documented rationale to report beyond 2 significant figures.

e. Rules for rounding the expanded uncertainty shall be defined by the laboratory.

f. The measurement result shall be reported using the same number of decimal places as the rounded expanded uncertainty unless a legal specification specifies how the measurement result is to be reported. Rules for rounding or truncating the measurement result shall be defined by the laboratory.

g. Laboratories shall not report the single largest measurement uncertainty for a group of analytes within a method or the largest measurement uncertainty for a single analyte across multiple methods.

h. For testing laboratories, if a significant bias is identified and the action taken is 4.2.6.1.1.1 b or c, this shall be clearly communicated.

4.3 Periodic evaluation of measurement uncertainty

The interval for review and recalculation of a method's MU shall be set by the laboratory. The interval and re-evaluation of measurement uncertainty will depend on, but not be limited to, the following factors:

a. Both Type A and Type B uncertainty components included in the calculation;

b. The frequency with which one of the components change;

c. The frequency with which the testing or calibration method is performed;

d. The magnitude of a change in a component in relationship to the calculated MU;

e. A change in the measurement process; and

f. Any laboratory administrative decision such as a set time interval.

Any recalculation of the measurement uncertainty shall meet all requirements of this standard.

ANNEX A
(Informative)

Concentration of ethanol in an ante-mortem blood specimen[9]

Test Method Information

Multiple analysts were trained and qualified to use the laboratory's method to determine the concentration of ethanol in ante-mortem blood specimens. All analysts use the same equipment for this test method. This includes a pipette diluter that delivers the specified sample volume together with a specified volume of aqueous internal standard.

The test method relies on gas chromatography with a flame ionization detector. Samples are introduced to the gas chromatograph via a headspace autosampler.

Calibrators are used to generate a calibration curve with each analytical batch. The calibrators are certified reference materials (CRMs) and span the reportable concentration range (e.g. 0.020 g/dL to 0.400 g/dL). The CRMs are not altered prior to use (i.e., not diluted). Constant variance (homoscedasticity) was observed across the concentration range. Method validation indicated that the proper calibration model was unweighted linear regression.

Measurement assurance is achieved through the use of Quality Control (QC) samples. These include a quantitative blood matrix control prepared by the laboratory at approximately 0.080 g/dL and unaltered CRMs at low, medium, and high concentrations (obtained from a different supplier than the CRMs used as calibrators). As with the CRMs used as calibrators, those used as QC samples are not altered prior to use.

Test specimens are analyzed in two separate batches. The average of the two measurement results is reported; however, the procedure requires that the individual measurements be no more than 5 % from the average or the analyses are repeated.

Calibrators, QC samples, and test samples are aliquoted in one instance using the same equipment.

Measurement Traceability

The traceability for this measurement process is established through the calibrators used to generate the calibration curve on the measuring system, as well as through the calibration of other equipment used in the measurement process.

[9] An evaluation of measurement uncertainty is specific to the measurement traceability that has been established for the measurement, the measurement assurance processes that are in place, the laboratory test method, the laboratory facility, etc. Therefore, the example that follows shall be evaluated and revised by each laboratory to take into consideration the elements that are specific to that laboratory.

All CRMs have been purchased from a Reference Material Producer that meets the *ANSI/ASB Standard 017, Standard Practices for Measurement Traceability in Forensic Toxicology*.

All external calibrations of measuring equipment are performed by calibration laboratories that meet the *ANSI/ASB Standard 017, Standard Practices for Measurement Traceability in Forensic Toxicology*. The pipette diluter has been and is routinely calibrated.

Measurement Assurance

The quantitative blood matrix control is prepared by the laboratory to a concentration of approximately 0.080 g/dL. It is made in a large batch, packaged, and stored in a manner to provide a long shelf-life for the control. The expected concentration is determined in-house through repeat measurements. Pre-defined criteria for acceptable performance are based on historical data across multiple lots from the last 2 years. To date, the laboratory has greater than 100 measurements made using this control since the method was validated.

The CRMs used for QC samples at low, medium, and high concentrations were purchased from a different supplier than the CRMs used as calibrators.

The QC samples are used to ensure validity of the test method across the concentration range. The CRM QC samples are also used to verify the calibration curve and to evaluate the method's bias on an ongoing basis.

Step 1 Specify the measurement process

As a written statement:

> *"The Concentration of Ethanol in Ante-Mortem Blood using [the validated laboratory procedure]"*

Step 2 Identify uncertainty components

The following list of *possible* contributors to the uncertainty in this method were identified by the laboratory:

<u>Analyst</u>

- Inter-analyst variation in sample preparation and measurements
- Training
- Experience

<u>Calibrators</u>

- CRM –uncertainty in the stated reference value
- Matrices of calibrators and test specimens

Quality Control Samples

- CRM – second source; uncertainty in the stated reference value
- Matrix control – stability

Internal Standard Preparation

- Components:
  - NaCl – reagent grade
  - n-propanol – reagent grade
  - Concentration – equipment used to prepare (balance, volumetric flask)

Preparation of aliquots of Calibrators, Quality Control Samples and Measurand

- Homogenization
  - Test Specimens – mixing
  - Matrix control – mixing

- Temperature
  - All calibrators, quality control samples and the test specimens are brought to room temperature
    - Variation in the time allowed to reach room temperature
    - Variation in room temperature at different times of year

- Pipette diluter
  - Volume of sample and volume of internal standard
    - Calibration uncertainty or laboratory specification to verify calibration status

- Headspace vials
  - Crimping action
  - Material of vial and stopper

- Time between replicate sampling of test specimens

Analysis

- Instrument parameter settings (e.g., oven temperature(s), gas flow, split ratio, aging of chromatographic column, autosampler syringe, autosampler precision, headspace equilibration time, headspace equilibration temperature, etc.)
- Interference from the matrix
- Interference from reagents
- Interference from other compounds
- Stability of sample(s) from preparation through analysis
- Instrument precision
- Systematic instrumental variation within an analytical batch

Data Processing

- Calibration model
- Integration parameters

- Processing algorithms

NOTE: This list of uncertainty components to be considered could also be compiled into a fishbone diagram or into any other format of the laboratory's choosing.

NOTE: A laboratory may identify different uncertainty components when an evaluation of their specific measurement process is performed.

Step 3 Quantify uncertainty components

The laboratory has existing data from the measurement process:

- The calibration model was determined during method validation and was shown through the use of residual plots to have constant variance across the linear range. Therefore, the laboratory is going to evaluate a single measurement uncertainty to represent the entire reportable concentration range.

- Each analytical batch does include one or more independently-prepared samples of the blood matrix quality control sample. This blood matrix QC sample is prepared to have an ethanol concentration of approximately 0.080 g/dL. All analysts have made measurements using this blood matrix QC sample (across multiple lots). To date, the laboratory has greater than 100 measurements of the blood matrix QC sample since validation.

- The laboratory also has data from three certified reference materials that were used as quality control samples. The ethanol concentration of the CRM QC samples spans the reportable concentration range. The primary use of the CRM QC samples is to evaluate bias in the measurement method, but these samples also provide additional evaluation of a number of uncertainty components.

Table 1 shows the individual uncertainty components and how they will be evaluated:

Table 1: Method of Evaluation of Uncertainty Components (Example 1)

Uncertainty Component	Method of Evaluation
Analysts	
Inter-analyst variation	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).
Training	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).
Experience	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).
Calibrators	

CRM – uncertainty in the stated reference value	*Type B Evaluation*
Matrices of calibrators and test specimens	Initially evaluated during method validation and determined to be insignificant, therefore not included in the uncertainty evaluation.
Quality Control Samples	
CRM – second source; uncertainty in the stated reference value	Primary use is to evaluate bias. The evaluation of bias will be done after the calculation of combined standard uncertainty.
Matrix control - stability	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).
Internal Standard Preparation	
Components: NaCl – reagent grade n-propanol – reagent grade	The measurement result will only be impacted by the volume of the internal standard added to each sample (i.e. variation due to pipette diluter).
Concentration- equipment used to prepare (balance, volumetric flask)	Procedural requirement to use the same lot of Internal Standard for all samples in an analytical batch. The measurement result will only be impacted by variation in the volume of the internal standard added to each sample (i.e. variation due to pipette diluter).
Preparation of aliquots of Calibrators, Quality Control Samples and Test Specimens	
Homogenization – mixing	Initially evaluated during method validation and determined to be significant, therefore controlled through the procedure administrative requirement for agreement of replicates *(Type B Evaluation).*
Temperature – all calibrators, quality control samples and the measurand are brought to room temperature Variation in the time allowed to reach room temperature Variation in room temperature at different times of year	Partially quantified in *Type A Evaluation* of process reproducibility data - blood matrix QC sample and partially through the procedure administrative requirement for agreement of replicates *(Type B Evaluation).*
Pipette diluter: Volume of sample, volume of internal standard and dilution Calibration uncertainty or laboratory specification to verify calibration status	*Type B Evaluation*
Pipette diluter: Variation in use by multiple staff	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).
Headspace vials: Crimping Material of stopper	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).

Time between replicate sampling of test item	Controlled through the procedure administrative requirement for agreement of replicates *(Type B Evaluation).*
Analysis	
Instrument parameter settings (e.g. oven temperature(s), gas flow, split ratios, aging of chromatographic column, autosampler syringe, autosampler precision, headspace equilibration time, headspace equilibration temperature, etc)	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).
Interference from the matrix	Duplicate listing of component – see Calibrators section above.
Interference from reagents	This component is not an uncertainty component but is a quality control concern. The laboratory analyzes a matrix blank that contains no analyte but does evaluate all reagents used in the analytical method. The laboratory procedure specifies acceptable criteria for this quality control sample.
Interference from other compounds	Initially evaluated during method validation and determined to be insignificant, therefore not included in the uncertainty evaluation.
Stability of sample(s) from preparation through analysis	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample) and through the procedure administrative requirement for agreement of replicates.
Instrument precision	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).
Systematic instrumental variation within an analytical batch	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample) and partially through the procedure administrative requirement for agreement of replicates *(Type B Evaluation).*
Data Processing	
Calibration model	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample and CRMs used as QC).
Integration parameters	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).
Processing algorithms	Adequately represented by the *Type A Evaluation* of process reproducibility data (Blood Matrix QC Sample).

***Type A Evaluation* of uncertainty components**

Measurement Process Reproducibility - Blood Matrix quality control sample

The number of observations of the blood matrix QC sample in this example is greater than 100. The statistic that will be calculated is the percent relative standard deviation.

To begin, the mean (average) and standard deviation of the blood matrix QC sample values will be calculated.[10]

The mean is calculated as:

$$\underline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{(x_1 + x_2 + x_3 + \cdots + x_n)}{n}$$

The mean of the reproducibility data in this example is 0.0798 g/dL.

The standard deviation is calculated as:

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}}$$

The standard deviation of the reproducibility data in this example is 0.0027 g/dL

Relative Standard Deviation (RSD) is calculated as:

$$RSD = \frac{s}{\bar{x}}$$

$$\%\, RSD = RSD \times 100\ \%$$

The %RSD of the reproducibility data in this example is:

$$RSD = \frac{0.0027\ g/dL}{0.0798\ g/dL} = 0.0341$$

$$\%\, RSD = 0.0341 \times 100 = 3.41\ \%$$

***Type B Evaluation* of uncertainty components**

[10] For the readability of the example, the display of digits used in all calculations was abbreviated. Best practice is to include and carry all digits through all calculations and only round the reported value and its uncertainty to the proper number of significant figures.

Interference from the matrix

The laboratory did evaluate matrix effects during method validation which resulted in the test method incorporating a dilution factor using the pipette diluter. Dilution of the sample, in combination with the procedural requirements to mix the test item minimizes matrix effects. The laboratory does acknowledge that it is impossible to evaluate all variations in test item matrix during method validation; therefore, the test method does include a blood matrix QC sample and a requirement for agreement between replicate samples to quantify the impact of matrix on the measurement.

NOTE: The laboratory procedural requirement for replicate agreement is an example of an administrative control that restricts variation in the measurement method. It is up to a laboratory to determine if such an administrative control will be used. The decision may be based on, but not limited to, knowledge of the measurement process, the impact of repeat analysis on cost and process efficiency, and the required expanded uncertainty. Measurement data may at times exceed the administrative limit, but may not be considered to be a statistical outlier, depending on its magnitude.

The laboratory procedure requires that two aliquots be taken from the homogenized test item and that the measured ethanol concentrations of the two aliquots must be within ±5 % of the average or the analysis is repeated.

The two uncertainty components – process reproducibility and interference from matrix – quantify a number of the same uncertainty components. The matrix control, over a longer period of time, holds the impact from the matrix constant while the effects from equipment, calibration, operators, and the laboratory environmental conditions vary. The replicate samples of the test item provide information on the test item matrix and a short–term evaluation of the effect from equipment, calibration, operators, and the laboratory environment.

Calibrators: Uncertainty in the reference value

The laboratory reviewed the calibration certificates from all CRMs used for the calibration curve. The greatest uncertainty is 0.000233 g/dL for the 0.010 g/dL CRM.

$$Relative\ uncertainty = \left(\frac{0.000233\ g/dL}{0.010\ g/dL}\right) * 100 = 2.33\ \%$$

Pipette Diluter

The laboratory has set internal criteria (±3 %) to ensure proper functioning of the pipette diluter. This is greater than the specifications for calibration used by the external calibration laboratory (±2 %). Additionally, the procedure to ensure proper functioning is performed quarterly compared to the external calibration which is performed annually.

Therefore, the laboratory criteria of ±3 % will be used to quantify variability for this uncertainty component.

Step 4 Convert quantities to standard uncertainties

The measurement unit

In this example, the estimated relative uncertainty is expressed as a percentage.

***Type A Evaluation* of uncertainty components**

Measurement Process Reproducibility data

Test specimens are sampled in duplicate, analyzed in two separate batches and the laboratory procedure for the reported ethanol concentration is to average the two results. Repeat measurements of the test specimens provide more information and more confidence that the reported result is the best estimate of the true value. The measurement process reproducibility data is based on single measurements of 0.08 g/dL blood matrix QC sample. Therefore, the %RSD of the mean is calculated by taking the %RSD of the measurement process and dividing by the square root of the number of measurements averaged to generate the reported ethanol concentration.

NOTE: If a single measurement result for the test specimens is selected to be reported (e.g., the lowest value), then the standard deviation of the mean calculation is not applicable.

NOTE: If the laboratory makes an equal number of multiple measurements of the quality control sample as it does of the test specimens and averages the results to evaluate the acceptability of the quality control sample, then the standard deviation of the mean calculation is not applicable.

The %RSD of the reproducibility data in this example is 3.41 %

The mathematical expression for %RSD of the mean:

$$\%RSD_{mean} = \frac{\%RSD}{\sqrt{n}}$$

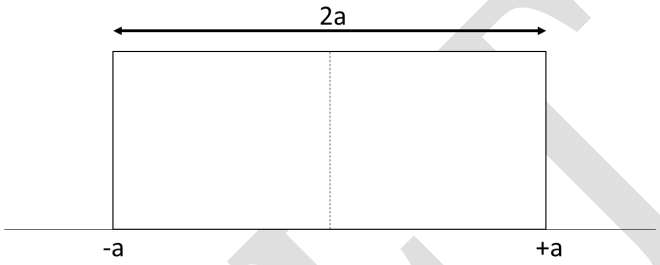
The %RSD of the mean of the reproducibility data in this example is:

$$\%RSD_{mean} = \frac{3.41\ \%}{\sqrt{2}} = 2.4101\ \%$$

Type B Evaluation **of uncertainty components**

Interference from the matrix

The laboratory procedure requires two samples to be taken from the homogenized test specimens and the ethanol concentration of the two aliquots to be within ±5 % of the average, or the analysis is repeated. This component is evaluated as a rectangular distribution:



Upper limit = +a
Lower limit = -a
Possible range of values = (+a) – (-a) = 2a

For a rectangular distribution, the standard uncertainty is calculated by:

$$Standard\ uncertainty = \frac{a}{\sqrt{3}}$$

The standard uncertainty for the interference from the matrix in this example is based on an outside limit of 5 %:

$$Standard\ uncertainty = \frac{5\ \%}{\sqrt{3}} = 2.8868\ \%$$

Calibrators: Uncertainty in the reference value

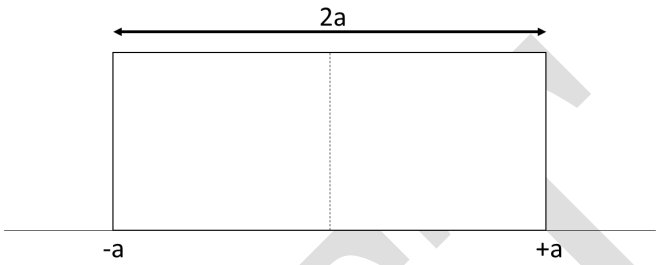
Based on the certificates from the CRMs used for calibrators in this method, the laboratory determined in Step 3 that the greatest relative uncertainty for the CRMs is 2.33 %.

The certificate indicates this expanded uncertainty assumes a normal distribution, a coverage factor of $k = 2$, and a coverage probability of approximately 95 %. The uncertainty on the calibration certificate will be divided by the coverage factor to arrive at a relative standard uncertainty.

$$Relative\ standard\ uncertainty = \left(\frac{2.33\ \%}{2}\right) = 1.165\ \%$$

Pipette Diluter

In Step 3, the laboratory determined that its in-house criteria of ±3 % will be used to quantify variability for this uncertainty component. This component is evaluated as a rectangular distribution:



Upper limit = +a
Lower limit = -a
Possible range of values = (+a) – (-a) = 2a

As explained above, for a rectangular distribution, the standard uncertainty is calculated by:

$$Standard\ uncertainty = \frac{a}{\sqrt{3}}$$

The standard uncertainty for the pipette diluter in this example is based on the outside limit of 3 %:

$$Standard\ uncertainty = \frac{3\ \%}{\sqrt{3}} = 1.7321\ \%$$

Step 5 Calculate the combined standard uncertainty

The evaluation will assume that the uncertainty components are independent or uncorrelated and that the measurement result is the sum of a series of components.

Care shall be taken if the measurement results lie over a range of values. In this scenario, the calibration model was determined during method validation and was shown through the use of residual plots to have constant variance across the linear range, so a single estimation of measurement uncertainty can be calculated for the entire concentration range.

$$u_c(y) = \sqrt{s^2_{reproducibility} + u^2_{matrix} + u^2_{CRMunc} + u^2_{pipettediluter}}$$

$$u_c(y) = \sqrt{2.4101^2_{reproducibility} + 2.8868^2_{matrix} + 1.165^2_{CRMunc} + 1.7321^2_{pipettediluter}}$$

$$u_c(y) = \sqrt{18.4992}$$

$$u_c(y) = 4.3011\,\%$$

Evaluation of bias

The laboratory views the monitoring of bias as a component of ensuring the validity of the test method and has incorporated three CRMs at a low, medium and high concentration as QC samples for the purpose of monitoring bias from unidentified sources on an ongoing basis.

The laboratory procedure requires each measured value for a CRM to be within 5 % of the reference value. The largest bias for any of the control levels (low, medium, and high) is less than the combined standard uncertainty. Although the bias is viewed as insignificant, the laboratory is choosing to include an additional component in the uncertainty evaluation that will address the uncertainty in the reference value of the CRM used for the evaluation of bias. Steps 3, 4, and 5 must be addressed for this additional uncertainty component.

Step 3 Quantify uncertainty components – bias component

The laboratory reviewed all of the calibration certificates from all CRMs used for the evaluation of bias. The greatest uncertainty is 0.0014 % for the 0.3 % CRM.

$$Reltative\ uncertainty = \left(\frac{0.0014\ \%}{0.3\ \%}\right) * 100 = 0.4667\ \%$$

Step 4 Convert quantities to standard uncertainties – bias component

The certificate indicates this expanded uncertainty assumes a normal distribution, a coverage factor of k = 2 and a coverage probability of approximately 95 %. The uncertainty on the calibration certificate will be divided by the coverage factor, 2, to arrive at a standard uncertainty.

$$\text{Relative standard uncertainty} = \left(\frac{0.4667\ \%}{2}\right) = 0.2334\ \%$$

Step 5 Calculate combined standard uncertainty – including bias component

The revised RSS calculation:

$$u_c(y) = \sqrt{s^2_{reproducibility} + u^2_{matrix} + u^2_{CRMunc} + u^2_{pipettediluter} + u^2_{CRMbias}}$$

$$u_c(y) = \sqrt{2.4101^2_{reproducibility} + 2.8868^2_{matrix} + 1.165^2_{CRMunc} + 1.7321^2_{pipettediluter} + 0.2334^2_{CRMbias}}$$

$$u_c(y) = \sqrt{18.5536}$$

$$u_c(y) = 4.3074\ \%$$

Step 6 Expand the combined standard uncertainty by coverage factor (*k*)

The data from the measurement process is assumed to follow a normal distribution. The laboratory has 101 measurements of the blood matrix quality control sample. Therefore, the laboratory assumes a lower bound on the effective degrees of freedom for the combined standard uncertainty of 100.

To expand the uncertainty to a 95.45 % coverage probability for this example, the coverage factor k = 2.025, from the Student's t-distribution table will be used.

$$U = 2.025 \times 4.3074 = 8.7225\ \%$$

NOTE: A laboratory can choose to increase the coverage probability.

Step 7 Evaluate the expanded uncertainty

The laboratory determined that the evaluation of uncertainty is fit-for-purpose based on the following considerations:

- Stakeholder interests
  Expanded uncertainty (8.7225 %) was below a requirement of 10 %.
- Legal requirements
  There were none.
- The relationship between the reported test value and the expanded MU
  Expanded uncertainty as a percentage across the analytical range ensures a consistent relationship.
- Established criteria including control limits for method
  The laboratory's control acceptance limits for the method are 5 % or 0.005 g/dL, whichever is larger. Considering the expanded uncertainty, the allowable control limits were modified to 8 % or 0.008 whichever is larger to minimize the occurrence of excessive QC failures.

Step 8 Report the uncertainty

The laboratory has established a procedure for rounding the expanded uncertainty. Following that procedure, the expanded uncertainty was rounded to two significant figures:

$$U = 8.7\ \%$$

For reporting measurement results with the rounded expanded uncertainty to the same number of decimal places:

> *"The concentration of ethanol in Item 1 was found to be 0.090 g/dL ± 0.008 g/dL at a coverage probability of 95.45%."*

ANNEX B

(Informative)

Concentration of amphetamine and methamphetamine in a whole blood specimen[11]

Test Method Information

The laboratory developed and validated a test method for quantitation of amphetamine and methamphetamine in whole blood, using liquid chromatography – tandem mass spectrometry (LC-MSMS). Multiple analysts were trained and qualified to use the laboratory's procedure. All analysts use the same equipment for this test method. Analytical results are normalized to internal standards added during the sample preparation process.

The method is calibrated using single replicates of whole blood fortified calibrators at 5 concentrations from 10 to 1000 ng/mL. The calibrators are prepared from a working stock solution that was made by diluting certified reference materials (CRMs). The working stock solution is fortified into whole blood with each batch. Method validation determined that the proper calibration model was quadratic regression model. Changing variance across the concentration range (heteroscedasticity) was observed across the concentration range.

The measurement results from single aliquots of a test specimens are reported.

Calibrators, QC samples, and test specimens are aliquoted at the same time using the same equipment.

Measurement Traceability

The traceability for this measurement process is established through the calibrators used to generate the calibration curve on the measuring system, as well as through the calibration of other equipment used in the measurement process.

All CRMs have been purchased from a Reference Material Producer that meets the *ANSI/ASB Standard 017, Standard Practices for Measurement Traceability in Forensic Toxicology*.

All external calibrations of measuring equipment are performed by calibration laboratories that meet the *ANSI/ASB Standard 017, Standard Practices for Measurement Traceability in Forensic Toxicology*. The pipettes and volumetric flasks have been and are routinely calibrated.

[11] An evaluation of measurement uncertainty is specific to the measurement traceability that has been established for the measurement, the measurement assurance processes that are in place, the laboratory test method, the laboratory facility, etc. Therefore, the example that follows shall be evaluated and revised by each laboratory to take into consideration the elements that are specific to that laboratory.

Measurement Assurance

The QC samples at low (30 ng/mL), medium (400 ng/mL), and high (800 ng/mL) concentrations are fortified into whole blood from a working stock solution by the laboratory with each batch. The working stock solution for the controls are prepared from CRMs purchased from a different supplier than the CRMs used as calibrators. The QC samples are used to ensure validity of the test method across the concentration range and to evaluate the method's bias on an ongoing basis.

The laboratory has 15 measurements made of the QC samples during validation for each concentration.

Since two analytes are involved in this measurement procedure, two separate uncertainty evaluations will be needed.

Step 1 Specify the measurement process

The measurement processes can be described in a written statement:

> *"The Concentration of Amphetamine in Whole Blood using [the validated laboratory procedure]"*
>
> *"The Concentration of Methamphetamine in Whole Blood using [the validated laboratory procedure]"*

Step 2 Identify uncertainty components

The following list of *possible* contributors to uncertainty in this method were identified by the laboratory:

<u>Analyst</u>

- Inter-analyst variation in sample preparation and measurements
- Training
- Experience

<u>Calibrators Preparation</u>

- Components:
  - Methanol – reagent grade
  - Concentration – equipment used to prepare (pipettes, volumetric flask)
- CRMs – uncertainty in the stated reference value

<u>Quality Controls Preparation</u>

- Components:
  - Methanol – reagent grade

- Concentration – equipment used to prepare (pipettes, volumetric flask)
- CRMs – uncertainty in the stated reference value

Internal Standard Preparation

- Components:
  - Methanol – reagent grade
  - Stable isotope labeled amphetamine and methamphetamine
- Impurities in the internal standard (unlabeled drug)
- Concentration – equipment used to prepare (pipettes, volumetric flask)

Preparation of aliquots of Calibrators, Quality Control Samples and Measurand

- Homogenization
  - Test Specimens – mixing

- Temperature
  - All calibrators, quality control samples and the test specimens are brought to room temperature
    - Variation in the time allowed to reach room temperature
    - Variation in room temperature at different times of year

- Pipettes
  - Volume of sample, calibrators, controls and internal standard
    - Calibration uncertainty or laboratory specification to verify calibration status

Analysis

- Instrument parameter settings (e.g., gradient, flow rate, aging of chromatographic column, autosampler syringe, autosampler precision, etc.)
- Interference from the matrix
- Interference from reagents
- Interference from other compounds
- Stability of sample(s) from preparation through analysis
- Instrument precision
- Systematic instrumental variation within an analytical batch
- Matrix effect (ionization suppression/enhancement)

Data Processing

- Calibration model
- Integration parameters
- Processing algorithms

NOTE: This list of uncertainty components to be considered could also be compiled into a fishbone diagram or into any other format of the laboratory's choosing.

NOTE: A laboratory may identify different uncertainty components when an evaluation of their specific measurement process is performed.

Step 3 Quantify uncertainty components

The laboratory has validation data from the measurement process:

- The calibration model was determined during method validation and was shown through the use of residual plots to have some heteroscedasticity (the variance was not constant across the linear range). Therefore, the laboratory is going to evaluate the measurement uncertainty using data from the control with the largest variance and apply it to the entire reportable concentration range.

- The QC samples at low (30 ng/mL), medium (400 ng/mL), and high (800 ng/mL) concentrations are fortified into whole blood from a working stock solution by the laboratory with each batch. All analysts have contributed to the 15 replicate measurements of the quality control samples at each concentration.

Table 2 shows the individual uncertainty components and how they will be evaluated:

Table 2: Method of Evaluation of Uncertainty Components (Example 2)

Uncertainty Component	Method of Evaluation
Analysts	
Inter-analyst variation	Adequately represented by the *Type A Evaluation* of process reproducibility data
Training	Adequately represented by the *Type A Evaluation* of process reproducibility data
Experience	Adequately represented by the *Type A Evaluation* of process reproducibility data
Calibrators Preparation	
Components: Methanol – reagent grade	Adequately represented by the *Type A Evaluation* of process reproducibility data
Concentration CRM – uncertainty in the stated reference value Equipment used to prepare (pipettes, volumetric flask)	*Type B Evaluation*
Quality Control Samples Preparation	
Components: Methanol – reagent grade	Adequately represented by the *Type A Evaluation* of process reproducibility data
Concentration CRM – uncertainty in the stated reference value Equipment used to prepare (pipettes, volumetric flask)	*Type B Evaluation (if necessary for bias)*
Internal Standard Preparation	

Components: Methanol – reagent grade	Adequately represented by the *Type A Evaluation* of process reproducibility data
Stable isotope labeled amphetamine and methamphetamine Impurities in the internal standard (unlabeled drug)	No influence Certificate of analysis from material provider indicates no impurity The measurement result will only be impacted by the volume of the internal standard added to each sample
Concentration- equipment used to prepare (pipettes, volumetric flask)	No influence Procedural requirement to use the same lot of Internal Standard for all samples in an analytical batch
Preparation of aliquots of Calibrators, Quality Control Samples and Test Specimens	
Homogenization – mixing	Demonstrated during method validation to be insignificant.
Temperature – all calibrators, quality controls and the measurand are brought to room temperature Variation in the time allowed to reach room temperature Variation in room temperature at different times of year	Adequately represented by the *Type A Evaluation* of process reproducibility data
Pipettes: Volume of sample, calibrators, quality controls , and internal standard Calibration uncertainty or laboratory specification to verify calibration status	Volume of internal standard adequately represented by the *Type A Evaluation* of process reproducibility data *Type B Evaluation* for volume of sample and calibrators (for controls only if necessary for bias)
Analysis	
Instrument parameter settings (e.g., gradient, flow rate, aging of chromatographic column, autosampler syringe, autosampler precision, etc.)	Adequately represented by the *Type A Evaluation* of process reproducibility data
Interference from the matrix	Matrix interference was evaluated during method validation and found to be insignificant for the matrix type allowed in this method.
Interference from reagents	This component is not an uncertainty component but is a quality control concern. The laboratory analyzes a matrix blank that contains no analyte that does evaluate all reagents used in the analytical method. The laboratory procedure specifies acceptable criteria for this quality control sample.
Interference from other compounds	Demonstrated lack of interference from other compounds during method validation. This component is not considered an uncertainty component.

Stability of sample(s) from preparation through analysis	Adequately represented by the *Type A Evaluation* of process reproducibility data
Instrument precision	Adequately represented by the *Type A Evaluation* of process reproducibility data
Systematic instrumental variation within an analytical batch	The positive controls are reinjected at the end of the batch and must meet predefined criteria
Data Processing	
Calibration model	Adequately represented by the *Type A Evaluation* of process reproducibility data
Integration parameters	Adequately represented by the *Type A Evaluation* of process reproducibility data
Processing algorithms	Adequately represented by the *Type A Evaluation* of process reproducibility data

***Type A Evaluation* of uncertainty components**

Measurement Process Reproducibility

The number of observations of each QC sample is 15. The statistic that will be calculated is the percent relative standard deviation.

Through validation, it was determined that the variance was not consistent across the calibration range. Therefore the reproducibility data from the multiple QC sample levels for either target compound may not be combined. The 400 ng/mL QC sample had the greatest variance and will be used for this evaluation.

To begin, the mean (average) and standard deviation of the control data will be calculated.

- The mean of the reproducibility data in this example is 404 ng/mL for amphetamine and 416 ng/mL for methamphetamine.

- The standard deviation of the reproducibility data in this example is 15.90 ng/mL for amphetamine and 12.01 ng/mL for methamphetamine.

  The %RSD of the reproducibility data in this example is 3.936 % for amphetamine and 2.888 % for methamphetamine.

***Type B Evaluation* of uncertainty components**

Calibrators Preparation

Uncertainty in the reference value

The laboratory reviewed the calibration certificates from all CRMs used for the preparation of the calibration working stock solutions. The largest uncertainty was 0.005 mg/mL for the 1.000 mg/mL amphetamine CRM and 0.006 mg/mL for the 1.000 mg/mL methamphetamine CRM.

$$Relative\ uncertainty\ of\ Amphetamine\ CRM = \left(\frac{0.005\ mg/mL}{1.000\ mg/mL}\right) * 100 = 0.5\ \%$$

$$Relative\ uncertainty\ of\ Methamphetamine\ CRM = \left(\frac{0.006\ mg/mL}{1.000\ mg/mL}\right) * 100 = 0.6\ \%$$

Uncertainty in pipettes

The laboratory reviewed the calibration certificates of all pipettes that may be used for preparation of the calibration working stock solution. The largest uncertainty was 0.74 µL for a 100µL pipette.

$$Relative\ uncertainty\ of\ Pipettes\ to\ Prep\ Cal\ Working\ Stock = \left(\frac{0.74\ \mu L}{100\ \mu L}\right) * 100 = 0.74\ \%$$

Uncertainty in volumetric flasks

The laboratory reviewed the calibration certificates of all volumetric flasks that may be used for preparation of the calibration working stock solution. The largest uncertainty was 0.0086 mL for a 25mL volumetric flask.

$$Relative\ uncertainty\ of\ Vol\ Flask\ to\ Prep\ Cal\ Working\ Stock = \left(\frac{0.0086\ mL}{25\ mL}\right) * 100 = 0.0344\ \%$$

Preparation of aliquots of Calibrators and Test Specimens

Uncertainty in pipettes

The laboratory reviewed the calibration certificates of all pipettes that may be used to fortify the calibrators from the working stock solution into whole blood. The method requires the same pipette to be used to add the internal standard to calibrators, controls, and test specimens. The largest uncertainty was 0.74 µL for a 100-µL pipette.

$$Relative\ uncertainty\ of\ Pipettes\ to\ Fortify\ Calibrator\ Samples = \left(\frac{0.74\ \mu L}{100\ \mu L}\right) * 100 = 0.754\ \%$$

$$Relative\ uncertainty\ of\ Pipettes\ to\ Delivery\ Internal\ Standard = \left(\frac{0.74\ \mu L}{100\ \mu L}\right) * 100 = 0.74\ \%$$

The laboratory reviewed the calibration certificates of all pipettes that may be used to aliquot the test item. The largest uncertainty was 6.9 µL for a 1000-µL pipette.

$$Relative\ uncertainty\ of\ Pipettes\ to\ Aliquot\ Test\ Samples = \left(\frac{6.9\ \mu L}{1000\ \mu L}\right) * 100 = 0.69\ \%$$

Step 4 Convert quantities to standard uncertainties

The measurement unit

In this example the estimated relative uncertainty is expressed as a percentage.

***Type A Evaluation* of uncertainty components**

Measurement Process Reproducibility data

The % RSD (s_r) of the reproducibility data in this example is 3.936 % for amphetamine and 2.888 % for methamphetamine.

***Type B Evaluation* of uncertainty components**

Calibrators Preparation

Uncertainty in the reference value

Based on the certificates from the CRMs used to prepare the calibrator working stock solutions in this method, the laboratory determined in Step 3 that the relative uncertainty is 0.5 % and 0.6 % for amphetamine and methamphetamine, respectively.

The certificates indicate the expanded uncertainties assume a normal distribution, a coverage factor of k = 2, and a coverage probability of approximately 95 %. The relative uncertainties will be divided by the coverage factor to arrive at relative standard uncertainties.

$$Relative\ standard\ uncertainty\ of\ Amphetamine\ CRM = \left(\frac{0.5\ \%}{2}\right) = 0.25\ \% = u_{CRM}$$

$$Relative\ standard\ uncertainty\ of\ Methamphetamine\ CRM = \left(\frac{0.6\ \%}{2}\right) = 0.30\ \% = u_{CRM}$$

Uncertainty in pipettes

In Step 3, the laboratory determined that among the pipettes used to prepare the working stock solutions, the largest relative uncertainty was 0.74 % for a 100-µL pipette.

The pipette's calibration certificate indicates this expanded uncertainty assumes a normal distribution, a coverage factor of k = 2.87, and a coverage probability of approximately 95 %. The relative uncertainty derived from the calibration certificate will be divided by the coverage factor, 2.87, to arrive at a relative standard uncertainty.

$$Relative\ standard\ uncertainty\ of\ Pipettes\ to\ Prep\ Calib\ Working\ Stock = \left(\frac{0.74\ \%}{2.87}\right) = 0.258\ \% = u_{CRMp}$$

Uncertainty in volumetric flasks

In Step 3, the laboratory determined that among the volumetric flasks used to prepare the working stock solutions, the largest relative uncertainty was 0.0344 % for a 25mL flask.

The volumetric flask's calibration certificate indicates this expanded uncertainty assumes a normal distribution, a coverage factor of k = 2, and a coverage probability of approximately 95 %. The relative uncertainty derived from the calibration certificate will be divided by the coverage factor, 2, to arrive at a relative standard uncertainty.

$$Relative\ standard\ uncertainty\ of\ Vol\ Flasks\ to\ Prep\ Calib\ Working\ Stock = \left(\frac{0.0344\ \%}{2}\right) = 0.172\ \% = u_{CRMv}$$

Preparation of aliquots of Calibrators and Test Specimens

Uncertainty in pipettes

In Step 3, the laboratory determined that among the pipettes used to fortify the calibrators from the working stock solution into whole blood, the largest relative uncertainty was 0.74 % for a 100µL pipette. The same pipette is used to fortify all samples with the internal standards.

The pipette's calibration certificate indicates this expanded uncertainty assumes a normal distribution, a coverage factor of k = 2.87 and a coverage probability of approximately 95 %. The uncertainty derived from the calibration certificate will be divided by the coverage factor to arrive at a relative standard uncertainty.

$$Relative\ standard\ uncertainty\ of\ Pipettes\ to\ Fortify\ Calibrator\ Samples = \left(\frac{0.74\ \%}{2.87}\right) = 0.258\ \% = u_{CALp}$$

$$Relative\ standard\ uncertainty\ of\ Pipette\ to\ Deliver\ Internal\ Standard = \left(\frac{0.74\ \%}{2.87}\right) = 0.258\ \% = u_{ISp}$$

The laboratory also determined in Step 3 that among the pipettes used to aliquot test specimens, the largest relative uncertainty was 0.69 % for a 1000-µL pipette.

The pipette's calibration certificate indicates this expanded uncertainty assumes a normal distribution, a coverage factor of k = 2.87, and a coverage probability of approximately 95 %. The uncertainty on the calibration certificate will be divided by the coverage factor, 2.87, to arrive at a relative standard uncertainty.

$$Relative\ standard\ uncertainty\ of\ Pipettes\ to\ Aliquot\ Test\ Samples = \left(\frac{0.69\ \%}{2.87}\right) = 0.24\ \% = u_{ITEMp}$$

Step 5 Calculate the combined standard uncertainty

The evaluation will assume that the uncertainty components are independent or uncorrelated and that the measurement result is the sum of a series of components.

For Amphetamine:

$$u_c(y) = \sqrt{3.936_r^2 + 0.25_{CRM}^2 + 0.258_{CRMp}^2 + 0.0172_{CRMv}^2 + 0.258_{CALp}^2 + 0.258_{ISp}^2 + 0.24_{ITEMp}^2}$$

$$u_c(y) = \sqrt{15.8122}$$

$$u_c(y) = 3.9765\ \%$$

For Methamphetamine:

$$u_c(y) = \sqrt{2.888_r^2 + 0.30_{CRM}^2 + 0.258_{CRMp}^2 + 0.0172_{CRMv}^2 + 0.258_{CALp}^2 + 0.258_{ISp}^2 + 0.24_{ITEMp}^2}$$

$$u_c(y) = \sqrt{8.6881}$$

$$u_c(y) = 2.9476\ \%$$

Evaluation of bias

The laboratory in this example views the monitoring of bias as a component of ensuring the validity of the test method and has incorporated three controls prepared from CRMs at a low, medium and high concentration as QC samples for the purpose of monitoring bias from unidentified sources on an ongoing basis.

The largest average bias for any of the control levels (low, medium and high) during validation was -2.4 % for amphetamine and 4.0 % for methamphetamine.

The bias for amphetamine is less than the combined standard uncertainty (3.9765 %) and is therefore insignificant. No additional component for the uncertainty of the CRM used to evaluate bias will be added.

The bias for methamphetamine is greater than the combined standard uncertainty (2.9476 %) and is therefore significant. Steps 3, 4, and 5 must be addressed for the methamphetamine bias component.

Step 3 Quantify uncertainty components – bias component

During validation the largest bias for methamphetamine was quantified to be 4.0 %.

Step 4 Convert quantities to standard uncertainties – bias component

The laboratory has chosen option 4.2.6.1.1.1 (c) to address the bias for methamphetamine that was determined to be significant. Following the guidance in section 3.2.5.5 of NIST SOP 29, the bias is treated as an uncorrected systematic error and the following equation applying a rectangular distribution is used to address the uncertainty of the difference component (u_d) in the MU evaluation:

$$u_d = \frac{bias}{\sqrt{3}} = \frac{4.0}{\sqrt{3}} = 2.3094$$

Step 5 Calculate combined standard uncertainty – including bias component

For Methamphetamine the updated root sum of the squares:

$$u_c(y) = \sqrt{2.888_r^2 + 0.30_{CRM}^2 + 0.37_{CRMp}^2 + 0.0172_{CRMv}^2 + 0.258_{CALp}^2 + 0.258_{ISp}^2 + 0.24_{ITEMp}^2 + 2.3094_d^2}$$

$$u_c(y) = \sqrt{14.0918}$$

$$u_c(y) = 3.7539\ \%$$

Step 6 Expand the combined standard uncertainty by coverage factor (*k*)

The data from the measurement process is assumed to follow a normal distribution. The laboratory has 15 measurements of the 400 ng/mL QC control. Therefore, the laboratory assumes that the effective degrees of freedom for the combined standard uncertainty cannot be lower than 14.

Refer to the Student's *t*-distribution table to determine the *k* factor.

To expand the uncertainty to a 95.45 % coverage probability for this example the coverage factor *k* = 2.20 will be used.

For Amphetamine:

$$U = 2.20 \times 3.9765 = 8.4079\ \%$$

For Methamphetamine:

$$U = 2.20 \times 3.7539 = 8.2586\ \%$$

Step 7 Evaluate the expanded uncertainty

The laboratory determined that the evaluation of uncertainty is fit-for-purpose based on the following considerations:

- Stakeholder interests
  There were none.
- Legal requirements
  There were none.
- The relationship between the reported test value and the expanded MU
  Expanded uncertainty as a percentage across the analytical range ensures a consistent relationship.
- Established criteria including control limits for method
  The laboratory's control limits for the method are 20 %. The allowable control limits were modified to 10 % for amphetamine and for methamphetamine to reflect the expanded uncertainty.

Step 8 Report the uncertainty

The laboratory has established a procedure for the process of rounding the expanded uncertainty. Following that procedure, the expanded uncertainty rounded to two significant figures:

For Amphetamine:

U = 8.4%

For Methamphetamine:

U = 8.3 %

For reporting measurement results with the rounded expanded uncertainties to the same number of decimal places:

"The concentration of amphetamine in Item 1 was found to be 90 ± 8 ng/mL at a coverage probability of 95.45%. The concentration of methamphetamine in Item 1 was found to be 143 ± 12 ng/mL at a coverage probability of 95.45%."

ANNEX C

(Informative)

Calibration of breath alcohol measuring instrumentation using long-term calibration data from a single instrument[12]

Calibration Method Information

The calibration of an individual breath alcohol instrument uses dry gas measurement standard data from the current calibration as well as historical calibration data for this single instrument over time. The calibration method uses measurement standards at multiple concentrations.

The calibration method does require each concentration of the dry gas measurement standards to be evaluated in triplicate. The method requires each of the triplicate measurements to be within 3 % or 0.003 g of ethanol/210 L of breath (g/210 L), whichever is greater, of the certified reference value of the measurement standard. Furthermore, the method requires that there shall be no greater than 0.003 g/210 L difference in all three measurements at each concentration.

Step 1 Specify the measurement process

Calibration of breath alcohol measuring instrumentation using long-term calibration data from a single instrument

Step 2 Identify uncertainty components

The following list of *possible* contributors to uncertainty in the calibration method were identified:

<u>Analyst</u>

- Inter-analyst variation in performing calibration
- Training
- Experience

<u>Breath Alcohol Measuring Instrument Being Calibrated</u>

- Variability of instrument over time

<u>Measurement Standards</u>

[12] An evaluation of measurement uncertainty is specific to the measurement traceability that has been established for the measurement, the measurement assurance processes that are in place, the laboratory calibration method, the laboratory facility, etc. Therefore, the example that follows shall be evaluated and revised by each laboratory to take into consideration the elements that are specific to that laboratory.

- Dry Gas Certified Reference Materials - uncertainty in the stated reference value

Environmental Conditions

- Barometric pressure
- Humidity
- Temperature

Varying Facilities/Location Change

- Instrument transport
- Power fluctuation

Data Processing

- Processing algorithms

Step 3 Quantify uncertainty components

Measurement standard data has been collected from use of this calibration method over time. All analysts have participated in acquiring the measurement standard data for this single breath alcohol measuring instrument. The laboratory has 51 measurements made using each measurement standard. The instrument has demonstrated constant variance across the concentration range of the measurement standards used in the calibration method. Because the 0.100 g/210 L measurement standard has the greatest observed variance of the measurement standards, it will be used to represent the process reproducibility data.

Table 1 shows the individual uncertainty components and how they will be evaluated:

Table 1: Method of Evaluation of Uncertainty Components (Example 3)

Uncertainty Component	Method of Evaluation
Analysts	
Inter-analyst variation	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard
Training	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard
Experience	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard
Breath Alcohol Measuring Instrument Being Calibrated	
Variability of instrument over time	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard
Measurement Standards	

CRM –uncertainty in the stated reference value	*Type B Evaluation*
Environmental Conditions	
Barometric pressure	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard
Humidity	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard
Temperature	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard
Varying Facilities/Locations	
Instrument transport	Not Applicable
Power fluctuations	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard.
Data Processing	
Processing algorithms	Adequately represented by *Type A Evaluation* of process reproducibility data – measurement standard

***Type A Evaluation* of uncertainty components**

Measurement Standard Reproducibility – 0.100 g/210 L Measurement Standard

The number of observations in this example is 51. The statistic that will be calculated is the standard deviation.

To begin, the mean (average) and standard deviation of the measurement data will be calculated.[13]

The mean is calculated as:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\bar{x} = \frac{(x_1 + x_2 + x_3 + \ldots x_n)}{n}$$

[13] For the readability of the example, the display of digits used in all calculations was abbreviated. Best practice is to include and carry all digits through all calculations and only round the reported value and its uncertainty to the proper number of significant figures.

The mean of the reproducibility data in this example = 0.0994 g/210 L

The standard deviation is calculated as:

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}}$$

The standard deviation of the reproducibility data in this example = 0.0012 g/210 L

***Type B Evaluation* of uncertainty components**

Certified Reference Materials

Based on the certificates from the CRMs, the laboratory determined in Step 3 that the greatest relative uncertainty for the CRM was 0.0018 g/210 L for the 0.100 g/210 L CRM.

The certificate indicates this expanded uncertainty assumes a normal distribution, a coverage factor of k = 2 and a coverage probability of approximately 95 %. The uncertainty on the calibration certificate will be divided by the coverage factor to arrive at a relative standard uncertainty.

$$Relative\ standard\ uncertainty = \left(\frac{0.0018\ g\ /210L}{2}\right) = 0.0009\,\frac{g}{210}L$$

Step 4 Convert quantities to standard uncertainties

The measurement unit: g of ethanol/210 L of breath (g/210 L)

***Type A Evaluation* of uncertainty components**

Measurement Standard Reproducibility – 0.100 g/210 L Measurement Standard

The standard deviation of the reproducibility data in this example is 0.0012 g/210 L.

- No additional conversion is necessary to reach a standard uncertainty.

Type B Evaluation* of uncertainty components*:

Certified Reference Materials

The CRM certificate indicates that the stated expanded uncertainty assumes a normal distribution, a coverage factor of k = 2 and a coverage probability of approximately 95 %.

- The uncertainty is stated to be 0.0018 g/210 L for the 0.100 g/210 L CRM.
- The uncertainty on the calibration certificate will be divided by the coverage factor, 2, to arrive at a standard uncertainty.
- 0.0018 g/210 L /2 = 0.0009 g/210 L for the standard uncertainty

Step 5 Calculate combined standard uncertainty

The evaluation will assume that the uncertainty components are independent or uncorrelated and that the measurement result is the sum of a series of components. The combined standard uncertainty was calculated.

$$u_c(y) = \sqrt{s^2_{reproducibility} + u^2_{CRMunc}}$$

$$u_c(y) = \sqrt{0.0012^2_{reproducibility} + 0.0009^2_{CRMunc}}$$

$$u_c(y) = \sqrt{0.0012^2_{reproducibility} + 0.0009^2_{CRMunc}}$$

$$u_c(y) = \sqrt{2.25x10^{-6}}$$

$$u_c(y) = 0.0015\ g/210L$$

Evaluation of Bias

In this example, bias is evaluated as part of instrument calibration.

The data for the 0.100 g/210 L measurement standard shows a difference of the average to reference value of 0.001 g/210 L. This value is less than the combined standard uncertainty and therefore, is insignificant. No additional component will be added to the measurement uncertainty evaluation.

Step 6 Expand the combined standard uncertainty by coverage factor (k)

The laboratory has 51 measurements of the measurement standard. Therefore, the laboratory assumes a lower bound on the effective degrees of freedom for the combined standard uncertainty of 50.

The data from the measurement process is assumed to follow a normal distribution; therefore, refer to the Student's *t*-distribution table to determine the k factor.

To expand the uncertainty to a 95.45 % coverage probability for this example the coverage factor k = 2.05 (n=50) will be used.

A laboratory can choose to increase the coverage probability.

$$k = 2.05$$

$$U = 2.05 \times 0.0015 = 0.00308\ g/210L$$

Step 7 Evaluate the expanded uncertainty

The calibration laboratory determined that the evaluation of uncertainty is fit-for-purpose.

The laboratory identified that the current method allows for a variance of 0.005 g/ 210L or 5%, whichever is greater, from a measurement standard known reference value. However, this is greater than the expanded uncertainty at 95.45 %. Left unchanged, a calibration could be reported that would have a bias that is significant. Therefore, the laboratory revised the method so that the variability allowed in any calibration must be equal to or less than the 0.003 g/ 210L or 3% whichever is greater.

Step 8 Report the uncertainty

The laboratory has established a procedure for the process of rounding the expanded uncertainty. Following that procedure, the expanded uncertainty is rounded to the third decimal place to equal the number of decimal places reported in the breath alcohol instrument display. The expanded uncertainty will be 0.003 g/210 L.

The certificate of calibration shall contain:

- 0.003 g/210 L, the combined expanded uncertainty, rounded to the third decimal place.
- k = 2.05, the coverage factor based on the student's t distribution.
- 95.45 %, the coverage probability

For reporting calibration results use the rounded expanded uncertainty to the same level of significance

0.100 g/210 L ± 0.003 g/210 L at a coverage probability of 95.45 % (k=2.05)."

ANNEX D

(Informative)

Calibration of breath alcohol measuring instruments using control data from the calibration method[14]

Calibration Method Information

A population of breath alcohol measuring instruments is calibrated using the same calibration method. The calibration method includes multiple measurement standards of varying concentrations and a control. The control data obtained is from a population of 100 breath alcohol measuring instruments that have all demonstrated constant variance across the measurement standard concentration levels. Three measurements of the 0.100 g of ethanol/210 L of breath (g/210 L) control is made during each instrument calibration. The current calibration as well as historical control data for the population of instruments over time was used in the calculation.

Step 1 Specify the measurement process

Calibration of breath alcohol measuring instruments using control data from the calibration method

Step 2 Identify uncertainty components

The following list of *possible* contributors to uncertainty in the calibration method were identified:

Analyst

- Inter-analyst variation in performing calibration
- Training
- Experience

Breath Alcohol Measuring Instrument Being Calibrated

- Population of 100 breath alcohol measuring instruments
- Variability of instrument over time

Measurement Standards

- Dry Gas Certified Reference Materials - uncertainty in the stated reference value

Calibration Method Control

[14] An evaluation of measurement uncertainty is specific to the measurement traceability that has been established for the measurement, the measurement assurance processes that are in place, the laboratory calibration method, the laboratory facility, etc. Therefore, the example that follows shall be evaluated and revised by each laboratory to take into consideration the elements that are specific to that laboratory.

- Dry Gas Certified Reference Material from a different manufacturer than that of the Measurement Standards - uncertainty in the stated reference value

Environmental Conditions

- Barometric pressure
- Humidity
- Temperature

Varying Facilities/Location Change

- Instrument transport
- Power fluctuations

Data Processing

- Processing algorithms

Step 3 Quantify uncertainty components

The calibration laboratory has existing data from the calibration method. Each instrument is evaluated, in triplicate, using a 0.100 g/210 L dry gas cylinder with measurement traceability as a calibration control. The calibration method requires the control to be within 3 % or 0.003 g/210 L (whichever is greater) of the certified reference value. Furthermore, there shall be no greater than 0.003 g/210 L difference in all three calibration control values.

Control data is collected on an on-going basis with all analysts contributing to the control data for the population of instruments.

Table 1 shows the individual uncertainty components and how they will be evaluated:

Table 1: Method of Evaluation of Uncertainty Components (Example 4)

Uncertainty Component	Method of Evaluation
Analysts	
Inter-analyst variation	Adequately represented by *Type A Evaluation* of process reproducibility data – control
Training	Adequately represented by *Type A Evaluation* of process reproducibility data – control
Experience	Adequately represented by *Type A Evaluation* of process reproducibility data – control
Breath Alcohol Measuring Instrument Being Calibrated	
Population of 100 breath alcohol measuring instruments	Adequately represented by *Type A Evaluation* of process reproducibility data – control

Variability of instrument over time	Adequately represented by *Type A Evaluation* of process reproducibility data – control
Measurement Standards	
CRM –uncertainty in the stated reference value	*Type B Evaluation*
Calibration Method Control	
CRM –uncertainty in the stated reference value	*Type B Evaluation*
Environmental Conditions	
Barometric pressure	Adequately represented by *Type A Evaluation* of process reproducibility data – control
Humidity	Adequately represented by *Type A Evaluation* of process reproducibility data – control
Temperature	Adequately represented by *Type A Evaluation* of process reproducibility data – control
Varying Facilities/Locations	
Instrument transport	Not Applicable
Power fluctuations	Adequately represented by *Type A Evaluation* of process reproducibility data – control.
Data Processing	
Processing algorithms	Adequately represented by *Type A Evaluation* of process reproducibility data – control

***Type A Evaluation* of uncertainty components**

Calibration Control Reproducibility – 0.100 g/210 L Calibration Control

The number of measurements of the control in this example is greater than 300.

The statistic that will be calculated is the standard deviation.

To begin, the mean (average) and standard deviation of the measurement data will be calculated.[15]

Mean

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

[15] For the readability of the example, the display of digits used in all calculations was abbreviated. Best practice is to include and carry all digits through all calculations and only round the reported value and its uncertainty to the proper number of significant figures.

$$\bar{x} = \frac{(x_1 + x_2 + x_3 + \ldots x_n)}{n}$$

The mean of the reproducibility data in this example = 0.0996 g/210 L

Standard Deviation

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}}$$

The standard deviation of the reproducibility data in this example = 0.0012 g/210 L

***Type B Evaluation* of uncertainty components**

Certified Reference Materials

The calibration laboratory reviewed the certificates of analysis from all dry gas cylinders. The greatest uncertainty is 0.0018 g/210 L for the 0.100 g/210 L CRM.

Step 4 Convert quantities to standard uncertainties

The measurement unit: g of ethanol/210 L of breath (g/210 L)

***Type A Evaluation* of uncertainty components**

Calibration Control Reproducibility – 0.100 g/210 L Calibration Control

The standard deviation of the reproducibility data in this example is 0.0012 g/210 L.

- No additional conversion is necessary to reach a standard uncertainty.

***Type B Evaluation* of uncertainty components**

Certified Reference Materials

The certificates of analysis state that the expanded uncertainty assumes a normal distribution, a coverage factor of $k = 2$ and a coverage probability of approximately 95 %.

- The greatest uncertainty is 0.0018 g/210 L.
- The uncertainty on the calibration certificate will be divided by the coverage factor, 2, to arrive at a standard uncertainty.
- 0.0018 g/210 L /2 = 0.0009 g/210 L for the standard uncertainty.

Step 5 Calculate combined standard uncertainty

The evaluation will assume that the uncertainty components are independent or uncorrelated and that the measurement result is the sum of a series of components. The combined standard uncertainty was calculated.

$$u_c(y) = \sqrt{s^2_{reproducibility} + u^2_{CRMunc}}$$

$$u_c(y) = \sqrt{0.0012^2_{reproducibility} + 0.0009^2_{CRMunc}}$$

$$u_c(y) = \sqrt{0.0012^2_{reproducibility} + 0.0009^2_{CRMunc}}$$

$$u_c(y) = \sqrt{2.25x10^{-6}}$$

$$u_c(y) = 0.0015\ g/210L$$

Evaluation of Bias

In this example, bias is evaluated as part of the instrument calibration. The calibration method requires the control to be within 3 % or 0.003 g/210 L (whichever is greater) of the certified reference value. Furthermore, there shall be no greater than 0.003 g/210 L difference in all three calibration control values.

The data for the 0.100 g/210 L calibration control shows a difference between the average and the reference value of 0.001 g/210 L. This value is less than the combined standard uncertainty and therefore, is insignificant. Although the bias is viewed as insignificant, the laboratory is choosing to include an additional component in the uncertainty evaluation. An uncertainty contributor equal to the uncertainty of the reference value of the calibration control used for the bias evaluation was added to the evaluation of measurement uncertainty.

Step 3 Quantify uncertainty components – bias component

The laboratory noted the difference of the average data for the 0.100 g/210 L calibration to be 0.001 g/210 L.

Step 4 Convert quantities to standard uncertainties – bias component

The standard uncertainty for the bias was 0.001 g/210 L.

Step 5 Calculate combined standard uncertainty – including bias component

The updated RSS calculation:

$$u_c(y) = \sqrt{s^2_{reproducibility} + u^2_{CRMunc} + u^2_{bias}}$$

$$u_c(y) = \sqrt{0.0012^2_{reproducibility} + 0.0009^2_{CRMunc} + 0.001^2_{bias}}$$

$$u_c(y) = \sqrt{0.0012^2_{reproducibility} + 0.0009^2_{CRMunc} + 0.001^2_{bias}}$$

$$u_c(y) = 0.0018\ g/210L$$

Step 6 Expand the combined standard uncertainty by coverage factor (k)

The data from the measurement process is assumed to follow a normal distribution.

The laboratory has 300 measurements of the calibration control. Refer to the Student's *t*-distribution table to determine the *k* factor.

To expand the uncertainty to a 95.45 % coverage probability for this example the coverage factor k = 2.0 will be used.

A laboratory can choose to increase the coverage probability.

$$k = 2.0$$

$$U = 2.0 \times 0.0018 = 0.0036\ g/210L$$

Step 7 Evaluate the expanded uncertainty

The calibration laboratory determined that the evaluation of uncertainty is fit-for-purpose.

Step 8 Report the uncertainty

The laboratory has established a procedure for the process of rounding the expanded uncertainty. Following that procedure, the expanded uncertainty rounded to the third decimal place. The expanded uncertainty will be 0.004 g/210 L.

The certificate of calibration shall contain:

- 0.004 g/210L, the combined expanded uncertainty, rounded to the third decimal place.
- k = 2.0, the coverage factor based on the student's t distribution.
- 95.45 %, the coverage probability

For reporting calibration results use the rounded expanded uncertainty to the same level of significance

0.100 g/210 L ± 0.004 g/210 L at a coverage probability of 95.45 % (k=2.0)."

ANNEX E

(informative)

Bibliography

1. ASCLD/LAB Policy on Measurement Uncertainty
   Note: Document can be obtained from anab@anab.org

2. ASCLD/LAB Guidance on the Estimation of Measurement Uncertainty – Overview
   Note: Document can be obtained from anab@anab.org

3. ASCLD/LAB Guidance on the Estimation of Measurement Uncertainty – Annex A Details on the NIST 8-Step Process
   Note: Document can be obtained from anab@anab.org

4. ASTM International E542-01 Standard Practice for Calibration of Laboratory Volumetric Apparatus available at www.astm.org

5. Eurachem Terminology in Analytical Measurement – Introduction to VIM 3 available at www.eurachem.org

6. The National Institute of Standards and Technology (NIST) definition of "internal measurement assurance program" available at www.nist.gov/traceability/index.cfm

7. Internal Organization for Standardization (ISO). ISO/IEC 9000:2015 Quality Management Systems-Fundamentals and Vocabulary (Geneva, Switzerland: ISO, 2000)

8. Internal Organization for Standardization (ISO). ISO/IEC 17000:2004 Conformity Assessment – Vocabulary and General Principles (Geneva, Switzerland: ISO, 2014)

9. ANSI/ASB Standard 036, Standard Practices for Method Validation in Forensic Toxicology. Available for download at XXX.

10. ANSI/ASB Standard 054, Standard for a Quality Control Program in Forensic Toxicology Laboratories. Available for download at XXX.

11. ANSI/ASB Standard 053, Standard for Report Content in Forensic Toxicology. Available for download at XXX.

12. ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories. Available for purchase at *https://webstore.ansi.org/*.